--DISCLOSURE OF THE INVENTION

Heretofore, many FO-extraction methods and apparatus have been proposed: time domain algorithm on the basis of interval measurement, frequency-domain method on the basis of spectrum, a method in which autocorrelation and harmonic sieve (sieve for extracting harmonic components) are used singly or in combination, and a biologically-motivated method. These methods and apparatus premise that a signal to be analyzed is a periodic signal from the viewpoint of mathematics. In each of these methods and apparatus, a value estimated on the basis of periodicity from the viewpoint of mathematics provides a correctly estimated FO value for a signal whose FO is constant over time. However, it is not clear whether conventional methods and apparatus can provide correctly estimated FO values in analysis of a real voice, where FO changes with time, or in analysis of complex sound in which the frequencies of sinusoidal-wave components deviate slightly from a harmonic relation.

In the proposed high-quality voice conversion system, conversion and re-synthesis of voice must be performed on the basis of accurate sound-source information of an original voice. Therefore, in order to improve this method, an FO-extraction method can rationally be applied to a signal whose

FO changes with time and a signal which includes non-harmonic components. Such an observation motivates the inventor to develop a new FO-extraction method and apparatus which produces an accurate FO locus with high temporal resolution by use of the instantaneous frequency of the fundamental component.

In the STRAIGHT method, an FO-extraction method based on instantaneous frequency has been developed and used on the assumption that a filtered signal containing a fundamental-wave component involves minimal AM modulation and FM modulation. The FO-extraction method used in the STRAIGHT method exhibited agreeable performance in an evaluation test which was performed while an EGG (Electro Glotto Graph) signal recorded simultaneously with voice was used as a reference signal. For example, in analysis of 100 sentences spoken by an adult female speaker, the error between FO obtained from voice and FO obtained from FGG became 20% or higher only in 1.4% of all analyzed frames. Further, in 53% of all analyzed frames, the F0 obtained from voice fell within 0.3% of the FO obtained from FGG. However, the above-described assumption of minimal AM and FM modulation is formulated ambiguously, and the formula is not effective mathematically. Further, this method involves a problem in that standard deviation of errors of FO regarding an adult male voice becomes about double that for an adult female voice.

The present invention provides a necessary mathematical base for enabling a new FO-extraction method and apparatus, which is an expansion of the above-described method. Detailed studies on partial differentiation of a function representing the relation between a filter center frequency and an output instantaneous frequency at a fixed point were key to providing a necessary mathematical base. Thus, the present invention leads to a new consistent FO/sound-source information extraction method and apparatus which utilizes a non-stationary aspect of the concept of instantaneous frequency.

An object of the present invention is to provide a method and apparatus for extracting sound-source information, which method enables the characteristics of fixed points of mapping from filter center frequency to output instantaneous frequency to be detected from instantaneous data, as a value which can be interpreted quantitatively.

[1] In a method and apparatus for extracting sound-source information by use of fixed points of mapping from frequency to instantaneous frequency, instantaneous frequency of each filter is partial-differentiated with respect to frequency to thereby obtain a first value; output of each filter is partial-differentiated with respect to frequency and then with respect to time to thereby obtain a second value; and proper weights are imparted to the first and second values and short-time

weighted integration with respect to time is performed to estimate a carrier-to-noise ratio of each filter, whereby a carrier-to-noise ratio is obtained, and an estimated value of evaluation value is obtained.

[2] In the method and apparatus for extracting sound-source information described in [1] above, on the basis of the evaluation value estimated by use of the carrier-to-noise ratio, a logarithm-frequency-axis analogous filter is used for selection of a fixed point corresponding to a fundamental frequency, and the fundamental frequency is extracted without advance information regarding the fundamental frequency.

[3] In the method and apparatus for extracting sound-source information described in [2] above, the logarithm-frequency axis analogous filter and a linear-frequency-axis analogous adapted chirp filter are used in combination in order to extract the fundamental frequency without advance information regarding the fundamental frequency and to improve the accuracy of the extracted fundamental frequency.--

IN THE CLAIMS

Add the following new claims:

--4 (New). An apparatus for extracting sound-source information by use of fixed points of mapping from frequency to instantaneous frequency, comprising:

means for performing partial differentiation of instantaneous frequency of each filter with respect to frequency to thereby obtain a first value;

means for performing partial differentiation of output of each filter with respect to frequency and then with respect to time to thereby obtain a second value; and

means for imparting proper weights to the first and second values and performing short-time weighted integration with respect to time to thereby estimate a carrier-to-noise ratio of each filter, whereby a carrier-to-noise ratio is obtained, and an estimated value of evaluation value is obtained.

5 (New). An apparatus for extracting sound-source information according to claim 4, further comprising a logarithm-frequency-axis analogous filter for selection of a fixed point corresponding to a fundamental frequency on the basis of the evaluation value estimated by use of the carrier-to-noise ratio, and means for extracting the fundamental frequency without advance information regarding the fundamental frequency.

6 (New). An apparatus for extracting sound-source information according to claim 5, wherein the logarithm-frequency-axis analogous filter and a linear-frequency-axis